



Enabling Grids for E-scienceE

# CMS Dashboard of Grid Activity

*Julia Andreeva, Juha Herrala, CERN  
LCG ARDA Project, EGEE NA4*

*EGEE User Forum  
Geneva, Switzerland  
March 1-3, 2006*

<http://arda.cern.ch>





- **ARDA is an LCG project**
  - Main objective is to enable analysis on the grid for LHC
  - Contribute to the distributed analysis activity of each LHC experiment
  - Build on top the new EGEE middleware (gLite)
    - *Use the grid software as it matures (exposed to all pre-releases)*
    - *Verify the components in an analysis environments*
    - *Provide **early and continuous** feedback*
- **Two main directions of the ARDA-CMS development**
  - **Dashboard** – monitoring system providing complete view of the Grid related and experiment specific information collected from the CMS distributed system
  - **Task Manager** and **Task Monitor** for supporting user analysis on the Grid



# The Dashboard project



- **The CMS Dashboard project aims to provide a single entry point for the monitoring data collected from the CMS distributed system.**
- **Dashboard development is a part of the grid integration project of the CMS experiment.**
  - Joined effort of the CMS experiment, the ARDA project and the MonAlisa team.
- **Dashboard will evolve from CMS-centric development into a joined project for CMS and ATLAS**
- **The activities have been concentrated on the job monitoring**
  - The objective is provide a complete view of how the jobs of a certain VO are doing on the Grid
  - Support for LCG/gLite and OSG middleware platforms.
- **Dashboard was used for the first time during the CMS SC3 (Service Challenge 3) in end of 2005.**



# Main principles



- **Provide a necessary level of interactivity**
  - *Not just precooked views, but possibility to dig in in case of troubles, possibility to ask detailed questions.*
- **Make the service pro-active**
  - *Foresee in future the possibility not only to collect and to expose information, but to analyze it and to generate alarms in case of evident problems*
- **Serve users with different level of requirements**
  - *Managers (computing projects managers, site managers, production managers) - CMS global views, site views*
  - *Users running their tasks on the grid - task views*
- **Follow input and feedback of the CMS community**
  - *What kind of information to collect*
  - *At what level to aggregate information*
  - *How to present collected information*

- **Quantities**

- How many jobs per site, per user, per submission tool, per data collection...
- Distribution over time

- **Usage of the resources**

- CPU, memory, IO rates
- Aggregated on different levels

- **Sharing of the resources**

- Between production and analysis
- Different analysis groups
- Individual users

- **How Grid is behaving**

- Success rate
- Failure reasons...

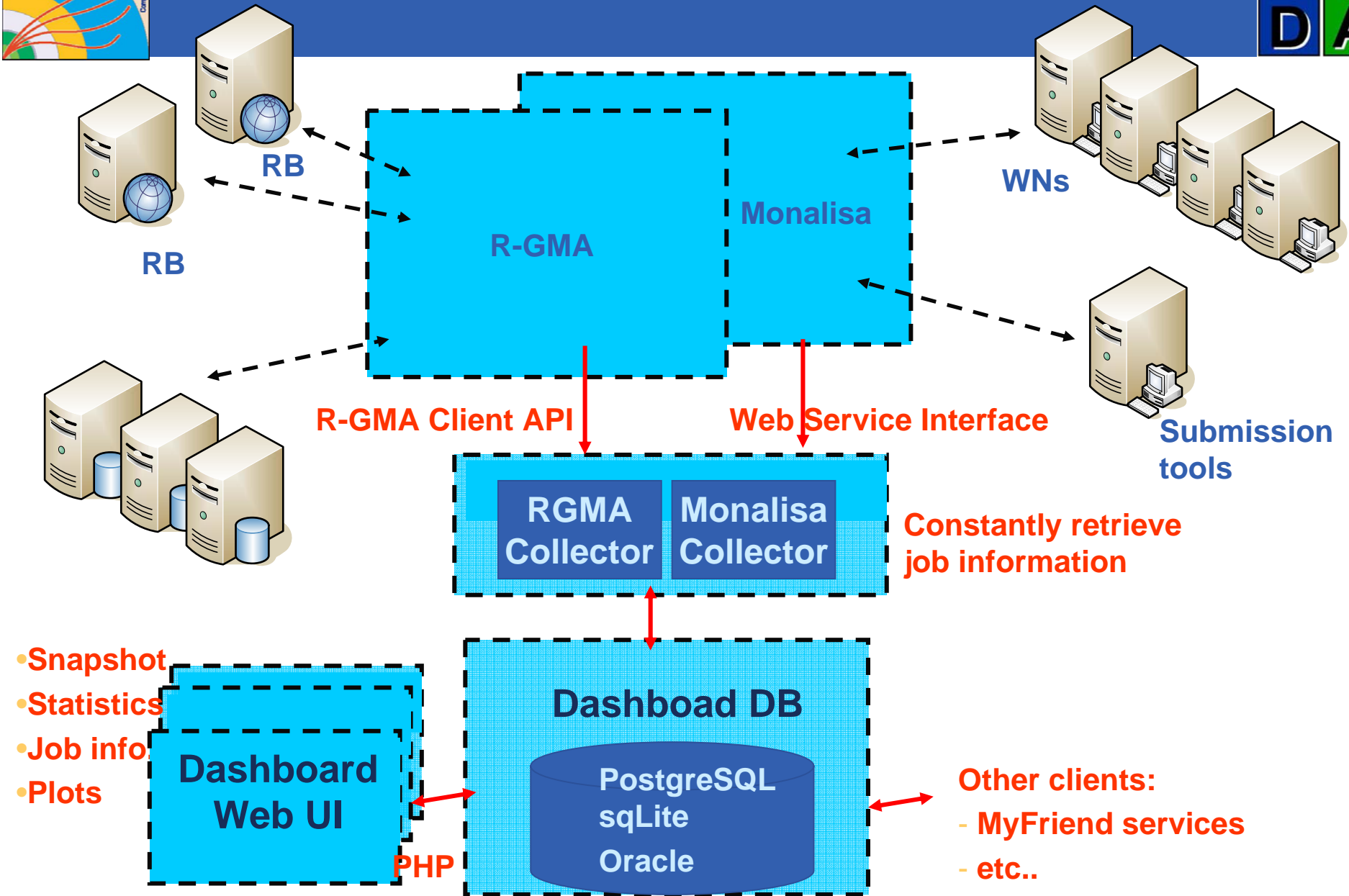
- **How CMS application is behaving**

- Success rate
- Failure reasons...

- **How CMS data is distributed over the sites**

- **Indication of the problems related to**

- *Pure Grid troubles (lack of stability, scalability, performance)*
- *Site configuration*
- *Data publishing*
- *Data access*
- *Software distribution*





# Sources of information



- **Currently two main sources of information are used**
- **R-GMA** for getting Grid related data (logging and book-keeping)
  - Under investigation possibility to use Gridlce and the LCG2 Real Time Monitor Resource Broker XML Files to complement the data.
- **MonAlisa** for **CMS-related data and system information**
  - CMS has a very positive experience for using MonAlisa for system monitoring.
  - In ARDA we started to use MonAlisa for job monitoring.
  - There is a plan to use MonAlisa to get batch system related information for the OSG sites
- **RGMA and MonAlisa are complementary in terms of the type of information they provide.**
  - Using two sources of information make possible crosschecking between the two.
  - Our experience shows that it allows to **increase the reliability of the system.**



# Use of the Dashboard during SC3



ASAP

Arda Support for cms Analysis Process



Conditions

user  site  submission tool

collection

sort by

type of view

Submitted After  , Submitted Before  , Status At   latest

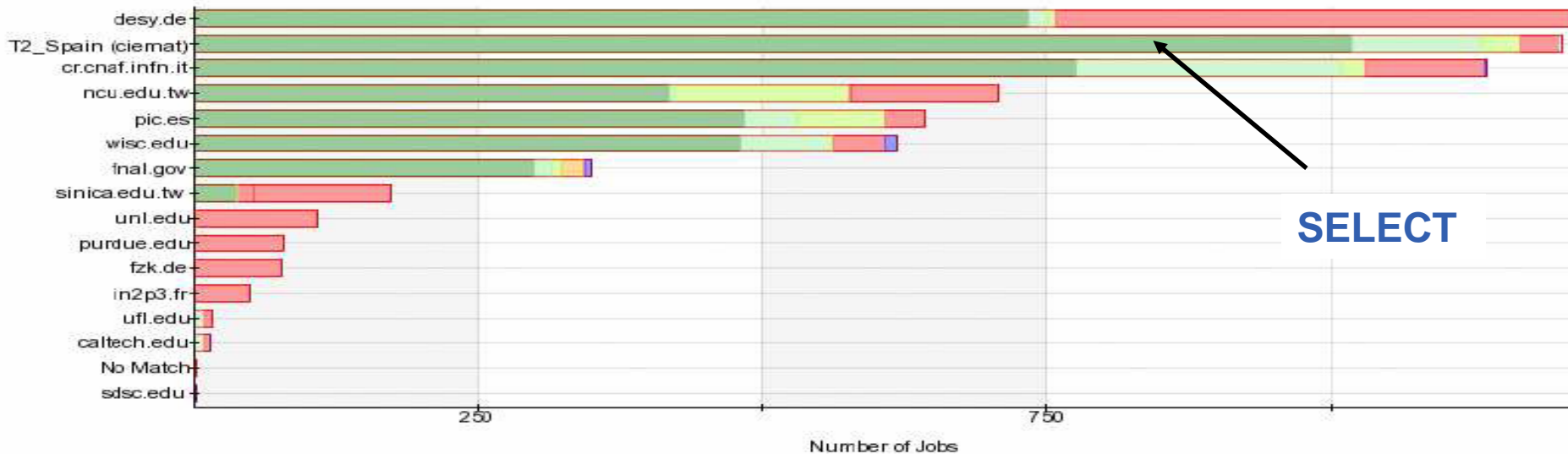
Note: time format is yyyy-mm-dd or yyyy-mm-dd hh:ii:ss  
Your current timezone is the same as [UTC+1 \(CERN\)](#), [click here](#) if you want to change it.

Jobs Status

show table:  plot:  [Note: How job status and success rate are calculated](#)

Site vs Number of Jobs for CRAB-v03-SC3

Site vs Number of Jobs for CRAB-v03-SC3



SELECT





# Use of the Dashboard during SC3

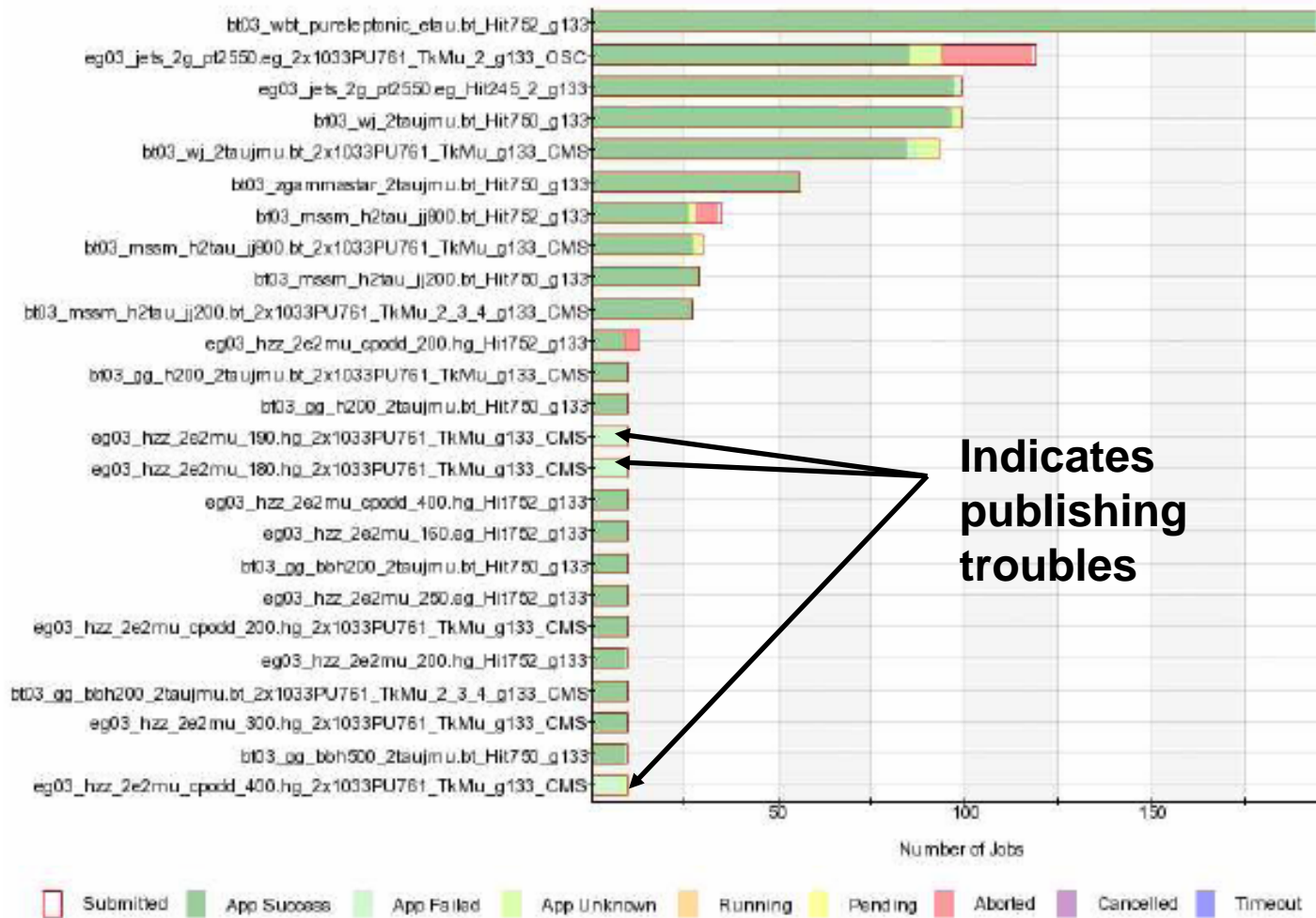


ASAP

Arda Support for cms Analysis Process



Collection vs Number of Jobs at T2\_Spain (ciemat) for CRAB-v03-SC3



Indicates publishing troubles

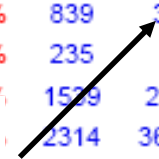


# Use of the Dashboard during SC3



site	Total Submitted	Grid Status of Submitted Jobs								Grid Success Rate	Application Exit Status of Finished Jobs		Application Success Rate	Processed Events	Overall Success Rate
		Pending	Running	Done	Timeout	Aborted	Assume Aborted	Cancelled	Unknown		Success	Failed			
wisc.edu	1207	-	-	1087	16	-	104	-	-	90.06%	998	85	91.81%	1012101	82.68%
unl.edu	443	-	-	188	143	-	112	-	-	42.44%	85	103	45.21%	127741	19.19%
purdue.edu	388	-	-	-	-	-	388	-	-	-	-	-	-	-	-
sdsc.edu	916	-	-	833	15	-	68	-	-	90.94%	22	809	2.64%	20935	2.4%
desy.de	2621	-	-	1978	29	501	113	-	-	75.47%	1920	28	97.07%	2026065	73.25%
fnal.gov	2064	501	-	1046	63	431	-	-	23	66.92%	839	37	80.21%	2167616	53.68%
fzk.de	580	1	-	240	-	160	179	-	-	41.45%	235	-	97.92%	238784	40.59%
pic.es	1909	-	-	1870	1	3	35	-	-	97.92%	1529	238	82.3%	1725232	80.62%
cr.cnaf.infn.it	9521	-	-	8960	12	104	445	-	-	91.51%	2314	3697	25.83%	2723948	24.3%
sinica.edu.tw	3070	42	-	954	15	1226	830	-	3	91.51%	446	456	46.75%	826242	14.73%
T2_Spain (ciemmat)	2414	-	-	2374	4	-	33	-	3	98.34%	2076	244	87.45%	2291446	86%
in2p3.fr	50	-	-	-	-	-	50	-	-	-	-	-	-	-	-
ncu.edu.tw	2435	7	-	1728	-	364	329	-	7	71.17%	1390	47	80.44%	1394492	57.25%
caltech.edu	108	-	-	67	1	-	40	-	-	62.04%	-	67	-	-	-
ufl.edu	31	-	-	24	-	-	7	-	-	77.42%	-	24	-	-	-
ba.infn.it	1544	-	-	946	-	598	-	-	-	61.27%	194	747	20.51%	195332	12.56%
UNKNOWN	40	-	-	4	-	15	21	-	-	10%	-	-	-	-	-
gridka.de	2	-	-	2	-	-	-	-	-	100%	2	-	100%	1995	100%
<b>Total :</b>	<b>29343</b>	<b>551</b>	<b>-</b>	<b>22301</b>	<b>299</b>	<b>3402</b>	<b>2754</b>	<b>-</b>	<b>36</b>	<b>77.46%</b>	<b>12060</b>	<b>6582</b>	<b>54.08%</b>	<b>14751929</b>	<b>41.89%</b>

SELECT





# Use of the Dashboard during SC3



**Conditions:**

User: ALL, Site: ALL, Job Submission Tool: CRAB-v03-SC3, task: ALL, collection: ALL  
 Submitted Time: 2005-12-05 ~ 2005-12-08, Status at 2006-02-01 17:11:04

Close

Scheduler JobId	site	Status Reason	JobMonitorId	TaskName	Events PerRun	Events Processed	Application Exit Status	Submitted at	Finished at
		user							
https://gdirb08.cern.ch:9000/U_qqUTSp4wI-zbAenq3pew	desy.de	retrieved	000046_131.169.223.107	SC3.051205.DESY.ExSi...	-	-	137	2005-12-05 15:59:33	2005-12-06 02:23:20
		output							
		san...							
		user							
https://gdirb01.cern.ch:9000/_BKI5QBW_9ZFvOKobLR2CA	desy.de	retrieved	000077_131.169.223.117	SC3.051205.DESY.ExRo...	-	55	137	2005-12-05 16:14:37	2005-12-06 15:14:06
		output							
		san...							
		user							
https://gdirb01.cern.ch:9000/jLDP-srA7nZU27pJSCfQ4Q	desy.de	retrieved	000047_131.169.223.114	SC3.051205.DESY.ExSi...	-	-	137	2005-12-05 16:34:55	2005-12-06 14:44:41
		output							
		san...							
		user							
https://gdirb01.cern.ch:9000/nCnCW9NKdixJEScfSBoMdQ	desy.de	retrieved	000074_131.169.223.92	SC3.051205.DESY.ExSi...	-	-	8	2005-12-05 16:44:44	2005-12-06 14:51:12
		output							
		san...							
		user							



# IO rates monitoring



## ASAP

Arda Support for cms Analysis Process



NOTE: This may involve long queries, please be patient!

Dynamic IO Rate

User

Site

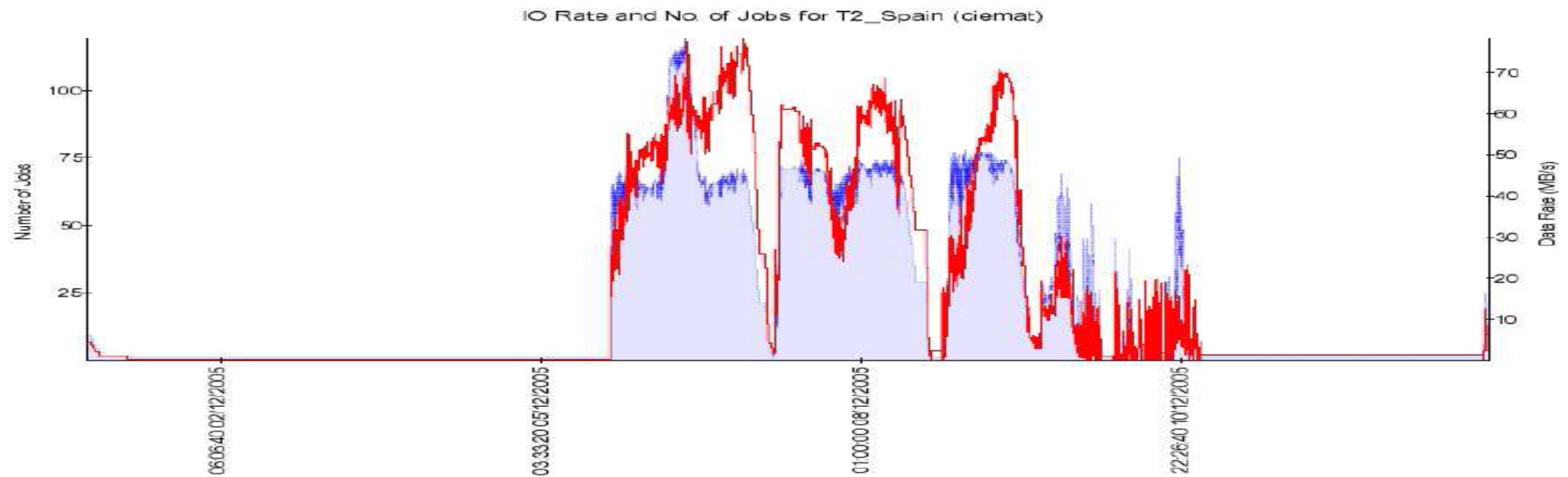
Job Submission Tool

Task

Collection

Start Date (d/m/y)        
HH:MM:SS:

End Date (d/m/y)        
HH:MM:SS:



- **ARDA analysis system is used by the CMS physicists for running their analysis tasks on the Grid and by CMS management to monitor the grid activities. Dashboard has got positive feedback from the community and will be integrated in the new CMS analysis system.**
- **First prototype of the Dashboard and its use during CMS service challenge demonstrated need of the experiment in the common monitoring tool combining Grid and experiment specific information. LHC experiments are running on several middleware platforms, so the single entry point of monitoring information has an additional value.**
- **Still a lot of work has to be done to pass from the prototype to a production quality system providing a needed level of reliability, scalability and performance.**